At page 14, line 24, delete "(*see, e.g.,* http://www.ncbi.nlm.nih.gov/Structure/RESEARCH/threading.html)".

At page 22, line 7, replace ")" with --), SEQ ID NO:1")--.

At page 37, line 7, replace "(mcrc@oligos.com)" with --(3112-A West Cuthbert Avenue, Midland, TX 79701)--.

At page 37, line 8, replace "(http://www.genco.com)" with --(1130 D Street, Suite #8, Ramona, CA 92065)--.

At page 37, line 8, replace "(www.expressgen.com)" with --(CTP Research Center, 2201 West Campbell Park Drive, Chicago IL 60612-3501), --.

At page 37, line 9, replace "(alameda, CA)" with --(1000 Atlantic Ave., Alameda, Ca)--.

At page 37, line 10, replace "(pkim@ccnet.com)" with --Research and Development Company--.

At page 37, line 10, replace " pro=ducts, Inc. (http://www.htibio.com)" with -- products, Inc.--.

At page 37, line 12, delete "Inc.," and insert --Inc.),--

After the Abstract, please add the accompanying sequence listing (1 page).

IN THE CLAIMS:

Please amend the claims by substituting the following claims for the corresponding previously pending claims of the same number(s):

1. A method of populating a data structure with a plurality of character strings, said method comprising:

i) encoding two or more biological molecules into character strings to provide a collection of two or more different initial character strings wherein each of said biological molecules comprises at least about 10 subunits;

ii) selecting at least two substrings from said character strings;

iii) concatenating said substrings to form one or more product strings about the same length as one or more of the initial character strings;

iv) adding the product strings to a collection of strings; and